



# SEQUENCE LISTING

<110> Koelsch, Gerald  
Tang, Jordan J. N.  
Hong, Lin  
Ghosh, Arun K.  
The Board of Trustees of the University of Illinois  
Oklahoma Medical Research Foundation

<120> Inhibitors of Memapsin 2 and Use Thereof

<130> 022266-000930US

<140> US 10/773,754

<141> 2004-02-06

<150> US 60/141,363

<151> 1999-06-28

<150> US 60/168,060

<151> 1999-11-30

<150> US 60/177,836

<151> 2000-01-25

<150> US 60/178,368

<151> 2000-01-27

<150> US 60/210,292

<151> 2000-06-08

<150> US 09/603,713

<151> 2000-06-27

<150> US 09/845,226

<151> 2001-04-30

<160> 39

<170> PatentIn Ver. 2.1

<210> 1

<211> 3252

<212> DNA

<213> Homo sapiens

<220>

<223> memapsin 2

<400> 1

```
gcgggagtgctgc tgcctgcccc cggcaccag caccggcatcc ggctgcccct gcgcagcggc 60
ctggggggcg cccccctggg gctgcggtctg ccccgggaga ccgacgaaga gcccgaggag 120
cccgcccgga ggggcagctt tgtggagatg gtggacaacc tgaggggcaa gtcggggcag 180
ggctactacg tggagatgac cgtgggcagc ccccgcgaga cgctcaacat cctgggtggat 240
acaggcagca gtaactttgc agtgggtgct gccccccacc ccttcctgca tcgctactac 300
cagaggcagc tgtccagcac ataccgggac ctccggaagg gtgtgtatgt gccctacacc 360
cagggaagtg ggggaaggga gctgggcacc gacctggtaa gcatcccca tggccccaac 420
gtcactgtgc gtgccaacat tgctgccatc actgaatcag acaagttctt catcaacggc 480
tccaactggg aaggcatcct ggggctggcc tatgctgaga ttgccaggcc tgacgactcc 540
ctggagcctt tctttgactc tctggtaaag cagaccacag ttcccaacct cttctcctg 600
cagctttgtg gtgctggctt cccctcaac cagtctgaag tgctggcctc tgctcgaggg 660
```

```

agcatgatca ttggagggtat cgaccactcg ctgtacacag gcagtctctg gtatacaccc 720
atccggcgagg agtgggtatta tgagggtgatc attgtgcggg tggagatcaa tggacaggat 780
ctgaaaatgg actgcaagga gtacaactat gacaagagca ttgtggacag tggcaccacc 840
aaccttcggt tgcccaagaa agtggtttgaa gctgcagtca aatccatcaa ggcagcctcc 900
tccacggaga agttccctga tggtttctgg ctaggagagc agctgggtgtg ctggcaagca 960
ggcaccaccc cttggaacat tttcccagtc atctcactct acctaattggg tgagggttacc 1020
aaccagtcct tccgcatcac catccttccg cagcaatacc tgcggccagt ggaagatgtg 1080
gccacgtccc aagacgactg ttacaagttt gccatctcac agtcatccac gggcactgtt 1140
atgggagctg ttatcatgga gggcttctac gttgtctttg atcgggccccg aaaacgaatt 1200
ggctttgctg tcagcgcttg ccatgtgcac gatgagttca ggacggcagc ggtggaaggc 1260
ccttttgtca ccttggacat ggaagactgt ggctacaaca ttccacagac agatgagtca 1320
accctcatga ccatagccta tgtcatggct gccatctgcg cctcttctcat gctgccactc 1380
tgccctatgg tgtgtcagtg gcgctgcctc cgctgcctgc gccagcagca tgatgacttt 1440
gctgatgaca tctccctgct gaagtggaga ggcccatggg cagaagatag agattccctt 1500
ggaccacacc tccgtggttc actttggtca caagtaggag acacagatgg cacctgtggc 1560
cagagcacct caggaccctc cccaccaccc aaatgcctct gccttgatgg agaaggaaaa 1620
ggctggcaag gtgggttcca gggactgtac ctgtaggaaa cagaaaagag aagaaagaag 1680
cactctgctg gcgggaatac tcttggtcac ctcaaattta agtcgggaaa ttctgctgct 1740
tgaaacttca gccctgaacc tttgtccacc attcctttta attotccaac ccaaagtatt 1800
cttcttttct tagtttcaga agtactggca tcacacgcag gttaccttgg cgtgtgtccc 1860
tgtggtaccc tggcagagaa gagaccaagc ttgtttccct gctggccaaa gtcagtagga 1920
gaggatgcac agtttgctat ttgctttaga gacagggact gtataaacia gcctaacatt 1980
ggtgcaaaga ttgcctcttg aattaaaaaa aaactagatt gactatttat acaaattggg 2040
gcggctggaa agaggagaag gagaggaggt acaaagacag ggaatagtgg gatcaaagct 2100
aggaaaggca gaaacacaac cactcaccag tcctagtttt agacctcatc tccaagatag 2160
catcccatct cagaagatgg gtgtgtgttt caatgttttc tttctgtgg ttgcagcctg 2220
acaaaagtg agatgggaag ggcttatcta gccaaagagc tcttttttag ctctcttaaa 2280
tgaagtggcc actaagaagt tccacttaac acatgaattt ctgccatatt aatttcattg 2340
tctctatctg aaccaccctt tattctacat atgataggca gcactgaaat atcctaacc 2400
cctaagctcc aggtgccctg tgggagagca actggactat agcagggtg ggctctgtct 2460
tcttggctat aggtcactc tttccccc aaatcttctct ggagctttgc agccaagggtg 2520
ctaaaaggaa taggtagag agctcttcta tctaatacctt aaaagcataa tgttgaacat 2580
tcattcaaca gctgatgccc tataaccctt gcctggattt ctctctatta ggctataaga 2640
agtagcaaga tctttacata attcagagtgt gtttcattgc ctctctaccc tctctaattg 2700
cccctccatt tatttgacta aagcatcrca cagtggcact agcattatac caagagtatg 2760
agaaatacag tgctttatgg ctctaacatt actgccttca gtatcaaggc tgcctggaga 2820
aaggatggca gcctcagggc ttccttatgt cctccaccac aagagctcct tgatgaagg 2880
catctttttc ccctatcctg ttcttcccct ccccgctcct aatggtacgt gggtagccag 2940
gctggttctt gggctaggtg gtggggacca agttcattac ctccctatca gttctagcat 3000
agtaaactac ggtaccagtg ttagtgggaa gagctgggtt ttcttagtat acccactgca 3060
tctactcct accctggtcaa cccgctgctt ccaggtatgg gacctgctaa gtgtggaatt 3120
acctgataag ggagagggaa atacaaggag ggcctctggt gttcctggcc tcagccagct 3180
gcccmaagc cataaacc aaamcaaga atactgagtc taaaaaaaa aaaaaaaaa 3240
aaaaaaaaaa aa 3252

```

<210> 2

<211> 488

<212> PRT

<213> Homo sapiens

<220>

<223> purified memapsin 2, aspartic proteinase 2 (ASP2)

<220>

<223> amino acids 28-48 are remnant putative propeptide residues





<210> 3  
 <211> 503  
 <212> PRT  
 <213> Homo sapiens  
  
 <220>  
 <223> pro-memapsin 2  
  
 <220>  
 <223> amino acids 1-13 are the T7 promoter  
  
 <220>  
 <223> amino acids 1-15 are vector-derived residues  
  
 <220>  
 <223> amino acids 16-64 are a putative pro-peptide  
  
 <220>  
 <223> amino acids 16-456 are pro-memapsin 2 T1  
  
 <220>  
 <223> amino acids 16-421 are promemapsin 2 T2  
  
 <400> 3  
 Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Ala  
   1                  5                  10                  15  
  
 Gly Val Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu  
           20                  25                  30  
  
 Arg Ser Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu  
           35                  40                  45  
  
 Thr Asp Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu  
   50                  55                  60  
  
 Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu  
   65                  70                  75                  80  
  
 Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr  
           85                  90                  95  
  
 Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His  
           100                  105                  110  
  
 Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys  
   115                  120                  125  
  
 Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly  
   130                  135                  140  
  
 Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala  
   145                  150                  155                  160  
  
 Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser  
           165                  170                  175  
  
 Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro  
   180                  185                  190

Asp	Asp	Ser	Leu	Glu	Pro	Phe	Phe	Asp	Ser	Leu	Val	Lys	Gln	Thr	His
		195					200					205			
Val	Pro	Asn	Leu	Phe	Ser	Leu	Gln	Leu	Cys	Gly	Ala	Gly	Phe	Pro	Leu
	210					215					220				
Asn	Gln	Ser	Glu	Val	Leu	Ala	Ser	Val	Gly	Gly	Ser	Met	Ile	Ile	Gly
225					230					235					240
Gly	Ile	Asp	His	Ser	Leu	Tyr	Thr	Gly	Ser	Leu	Trp	Tyr	Thr	Pro	Ile
				245					250					255	
Arg	Arg	Glu	Trp	Tyr	Tyr	Glu	Val	Ile	Ile	Val	Arg	Val	Glu	Ile	Asn
			260					265					270		
Gly	Gln	Asp	Leu	Lys	Met	Asp	Cys	Lys	Glu	Tyr	Asn	Tyr	Asp	Lys	Ser
		275					280					285			
Ile	Val	Asp	Ser	Gly	Thr	Thr	Asn	Leu	Arg	Leu	Pro	Lys	Lys	Val	Phe
	290					295					300				
Glu	Ala	Ala	Val	Lys	Ser	Ile	Lys	Ala	Ala	Ser	Ser	Thr	Glu	Lys	Phe
305					310					315					320
Pro	Asp	Gly	Phe	Trp	Leu	Gly	Glu	Gln	Leu	Val	Cys	Trp	Gln	Ala	Gly
				325					330					335	
Thr	Thr	Pro	Trp	Asn	Ile	Phe	Pro	Val	Ile	Ser	Leu	Tyr	Leu	Met	Gly
			340					345					350		
Glu	Val	Thr	Asn	Gln	Ser	Phe	Arg	Ile	Thr	Ile	Leu	Pro	Gln	Gln	Tyr
		355					360					365			
Leu	Arg	Pro	Val	Glu	Asp	Val	Ala	Thr	Ser	Gln	Asp	Asp	Cys	Tyr	Lys
	370					375					380				
Phe	Ala	Ile	Ser	Gln	Ser	Ser	Thr	Gly	Thr	Val	Met	Gly	Ala	Val	Ile
385					390					395					400
Met	Glu	Gly	Phe	Tyr	Val	Val	Phe	Asp	Arg	Ala	Arg	Lys	Arg	Ile	Gly
				405					410					415	
Phe	Ala	Val	Ser	Ala	Cys	His	Val	His	Asp	Glu	Phe	Arg	Thr	Ala	Ala
			420					425					430		
Val	Glu	Gly	Pro	Phe	Val	Thr	Leu	Asp	Met	Glu	Asp	Cys	Gly	Tyr	Asn
		435					440					445			
Ile	Pro	Gln	Thr	Asp	Glu	Ser	Thr	Leu	Met	Thr	Ile	Ala	Tyr	Val	Met
	450					455					460				
Ala	Ala	Ile	Cys	Ala	Leu	Phe	Met	Leu	Pro	Leu	Cys	Leu	Met	Val	Cys
465					470					475					480
Gln	Trp	Arg	Cys	Leu	Arg	Cys	Leu	Arg	Gln	Gln	His	Asp	Asp	Phe	Ala
				485					490					495	
Asp	Asp	Ile	Ser	Leu	Leu	Lys									
			500												

<210> 4  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:synthetic  
peptide derived from amyloid precursor protein  
(APP) beta-secretase site

<400> 4  
Ser Glu Val Lys Met Asp Ala Glu Phe Arg  
1 5 10

<210> 5  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:synthetic  
peptide derived from amyloid precursor protein  
(APP) beta-secretase site containing the "Swedish  
mutation"

<400> 5  
Ser Glu Val Asn Leu Asp Ala Glu Phe Arg  
1 5 10

<210> 6  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:synthetic  
peptide derived from the processing site of  
presenilin 1

<400> 6  
Ser Val Asn Met Ala Glu Gly Asp  
1 5

<210> 7  
<211> 12  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:synthetic  
peptide derived from amyloid precursor protein  
(APP) gamma-secretase site

<400> 7  
Lys Gly Gly Val Val Ile Ala Thr Val Ile Val Lys  
1 5 10

<210> 8  
<211> 3  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:signature  
aspartic protease sequence motif

<220>  
<221> MOD\_RES  
<222> (2)  
<223> Xaa = Thr or Ser

<400> 8  
Asp Xaa Gly  
1

<210> 9  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:synthetic  
peptide PS1-gamma in vivo beta-processing site of  
human presenilin 1, human memapsin 2 cleavage  
substrate

<400> 9  
Leu Val Asn Met Ala Glu Gly Asp  
1 5

<210> 10  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:PCR  
amplification primer NHASPF1

<400> 10  
ggtaagcatc ccccatggcc ccaacgtc

28

<210> 11  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:PCR  
amplification primer NHASPR1

<400> 11  
gacgttgggg ccatggggga tgcttacc

28



<210> 12  
 <211> 34  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:PCR  
         amplification primer NHASPF2  
  
 <400> 12  
 acgttggtctt tgatcgggcc cgaaaacgaa ttgg 34  
  
 <210> 13  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:PCR  
         amplification primer NHASPR2  
  
 <400> 13  
 ccaattcggtt ttctgggccc atcaaagaca acg 33  
  
 <210> 14  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:PCR  
         amplification adapter primer AP1  
  
 <400> 14  
 ccataccta acgactcact atagggc 27  
  
 <210> 15  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:PCR  
         amplification nested adapter primer AP2  
  
 <400> 15  
 actcactata gggctcgagc ggc 23  
  
 <210> 16  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:outside primer  
         GT10FWD

<400> 16  
 cttttgagca agttcagcct ggttaa 26

<210> 17  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:outside primer  
 GT10REV

<400> 17  
 gaggtggcctt atgagtattt cttccagggt a 31

<210> 18  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:outside primer  
 GT11FWD

<400> 18  
 tggcgacgac tcctggagcc cg 22

<210> 19  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:outside primer  
 GT11REV

<400> 19  
 tgacaccaga ccaactggta atgg 24

<210> 20  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:amplification  
 primer PASPN1

<400> 20  
 catatggcgg gagtgctgcc tgccac 27

<210> 21  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:amplification  
primer NHASPC1

<400> 21  
ggatcctcac ttcagcaggg agatgtcatc agcaaagt

38

<210> 22  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:oxidized bovine  
insulin beta chain hydrolytic site #5

<220>  
<221> MOD\_RES  
<222> (3)  
<223> Xaa = cysteic acid

<400> 22  
His Leu Xaa Gly Ser His Leu Val  
1 5

<210> 23  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:oxidized bovine  
insulin beta chain hydrolytic site #6

<220>  
<221> MOD\_RES  
<222> (1)  
<223> Xaa = cysteic acid

<400> 23  
Xaa Gly Glu Arg Gly Phe Phe Tyr  
1 5

<210> 24  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:synthetic  
peptide NCH-gamma hydrolytic site #7

<400> 24  
Val Gly Ser Gly Val  
1 5

<210> 25  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:synthetic  
       peptide NCH-gamma hydrolytic site #8  
  
 <400> 25  
 Val Gly Ser Gly Val Leu Leu  
   1                  5  
  
 <210> 26  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:synthetic  
       peptide NCH-gamma hydrolytic site #9  
  
 <400> 26  
 Gly Val Leu Leu Ser Arg Lys  
   1                  5  
  
 <210> 27  
 <211> 7  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:synthetic  
       memapsin 2 inhibitor substrate analogue OM99-1  
  
 <220>  
 <221> MOD\_RES  
 <222> (3)..(4)  
 <223> Xaa at positions 3 and 4 represent Leu and Ala  
       with the peptide bond substituted by a  
       transition-state isostere hydroxyethylene  
       (-CH(OH)-CH<sub>2</sub>-) group  
  
 <400> 27  
 Val Asn Xaa Xaa Ala Glu Phe  
   1                  5  
  
 <210> 28  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:octapeptide  
       upon which synthetic memapsin 2 inhibitor  
       substrate analogue OM99-2 is based

<400> 28  
 Glu Val Asn Leu Ala Ala Glu Phe  
 1 5

<210> 29  
 <211> 4  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:retained  
 structure of synthetic memapsin 2 inhibitor  
 substrate analogue OM99-1 with outside subsites  
 P4, P3, P3' and P4' discarded

<220>  
 <221> MOD\_RES  
 <222> (2)..(3)  
 <223> Xaa at positions 2 and 3 represent Leu and Ala  
 with the peptide bond substituted by a  
 transition-state isostere hydroxyethylene  
 (-CH(OH)-CH<sub>2</sub>-) group

<400> 29  
 Asn Xaa Xaa Ala  
 1

<210> 30  
 <211> 10  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:synthetic  
 peptide NCH-gamma

<400> 30  
 Val Gly Ser Gly Val Leu Leu Ser Arg Lys  
 1 5 10

<210> 31  
 <211> 326  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <223> pepsin

<220>  
 <223> amino acids 2-5, 6-9, 13-20, 25-32, 65-67, 69-74,  
 79-87, 89-91, 99-106, 119-122, 150-154, 164-167,  
 180-183, 191-194, 196-199, 201-204, 210-214,  
 221-223, 258-262, 265-269, and 275-278 are beta  
 strands

<220>  
 <223> amino acids 281-284, 286-288, 298-301, 310-315,  
 and 319-324 are beta strands

<220>

<223> amino acids 48-51, 111-114, 136-142, 225-234,  
249-254, 271-274, and 303-306 are helices

<220>

<223> amino acids 12-13, 30, 32, 34-35, 73-77, 111, 117,  
120, 189, 213, 215, 217-220, 287, 289, 291, 298,  
and 300 are residues in contact with pepstatin.

<400> 31

Val	Asp	Glu	Gln	Pro	Leu	Glu	Asn	Tyr	Leu	Asp	Met	Glu	Tyr	Phe	Gly
1				5					10					15	

Thr	Ile	Gly	Ile	Gly	Thr	Pro	Ala	Gln	Asp	Phe	Thr	Val	Val	Phe	Asp
			20					25					30		

Thr	Gly	Ser	Ser	Asn	Leu	Trp	Val	Pro	Ser	Val	Tyr	Cys	Ser	Ser	Leu
		35					40					45			

Ala	Cys	Thr	Asn	His	Asn	Arg	Phe	Asn	Pro	Glu	Asp	Ser	Ser	Thr	Tyr
	50					55					60				

Gln	Ser	Thr	Ser	Glu	Thr	Val	Ser	Ile	Thr	Tyr	Gly	Thr	Gly	Ser	Met
65					70					75					80

Thr	Gly	Ile	Leu	Gly	Tyr	Asp	Thr	Val	Gln	Val	Gly	Gly	Ile	Ser	Asp
				85					90					95	

Thr	Asn	Gln	Ile	Phe	Gly	Leu	Ser	Glu	Thr	Glu	Pro	Gly	Ser	Phe	Leu
			100					105					110		

Tyr	Tyr	Ala	Pro	Phe	Asp	Gly	Ile	Leu	Gly	Leu	Ala	Tyr	Pro	Ser	Ile
		115					120						125		

Ser	Ser	Ser	Gly	Ala	Thr	Pro	Val	Phe	Asp	Asn	Ile	Trp	Asn	Gln	Gly
	130					135					140				

Leu	Val	Ser	Gln	Asp	Leu	Phe	Ser	Val	Tyr	Leu	Ser	Ala	Asp	Asp	Gln
145					150					155					160

Ser	Gly	Ser	Val	Val	Ile	Phe	Gly	Gly	Ile	Asp	Ser	Ser	Tyr	Tyr	Thr
				165					170					175	

Gly	Ser	Leu	Asn	Trp	Val	Pro	Val	Thr	Val	Glu	Gly	Tyr	Trp	Gln	Ile
			180					185					190		

Thr	Val	Asp	Ser	Ile	Thr	Met	Asn	Gly	Glu	Ala	Ile	Ala	Cys	Ala	Glu
		195					200					205			

Gly	Cys	Gln	Ala	Ile	Val	Asp	Thr	Gly	Thr	Ser	Leu	Leu	Thr	Gly	Pro
	210					215					220				

Thr	Ser	Pro	Ile	Ala	Asn	Ile	Gln	Ser	Asp	Ile	Gly	Ala	Ser	Glu	Asn
225					230					235					240

Ser	Asp	Gly	Asp	Met	Val	Val	Ser	Cys	Ser	Ala	Ile	Ser	Ser	Leu	Pro
				245					250					255	

Asp	Ile	Val	Phe	Thr	Ile	Asn	Gly	Val	Gln	Tyr	Pro	Val	Pro	Pro	Ser
			260					265					270		

Ala Tyr Ile Leu Gln Ser Glu Gly Ser Cys Ile Ser Gly Phe Gln Gly  
275 280 285

Met Asn Leu Pro Thr Glu Ser Gly Glu Leu Trp Ile Leu Gly Asp Val  
290 295 300

Phe Ile Arg Gln Tyr Phe Thr Val Phe Asp Arg Ala Asn Asn Gln Val  
305 310 315 320

Gly Leu Ala Pro Val Ala  
325

<210> 32  
<211> 19  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:beginning of  
promemapsin 2 T1 and T2 construct

<400> 32  
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Ala  
1 5 10 15

Gly Val Leu

<210> 33  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:ending of  
promemapsin 2 T1 construct

<400> 33  
Gln Thr Asp Glu Ser Thr  
1 5

<210> 34  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:ending of  
promemapsin 2 T2 construct

<400> 34  
Gly Phe Ala Val Ser Ala  
1 5

<210> 35  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence:synthetic  
 memapsin 2 inhibitor substrate analogue OM99-2  
  
 <220>  
 <221> MOD\_RES  
 <222> (4)..(5)  
 <223> Xaa at positions 4 and 5 represent Leu and Ala  
 with the peptide bond substituted by a  
 transition-state isostere hydroxyethylene  
 (-CH(OH)-CH<sub>2</sub>-) group  
  
 <400> 35  
 Glu Val Asn Xaa Xaa Ala Glu Phe  
 1 5  
  
 <210> 36  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:pro-memapsin 2  
 hydrolytic site #1 (aa 12-18 of SEQ ID NO:3)  
  
 <400> 36  
 Arg Gly Ser Met Ala Gly Val Leu  
 1 5  
  
 <210> 37  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:pro-memapsin 2  
 hydrolytic site #2 (aa 23-30 of SEQ ID NO:3)  
  
 <400> 37  
 Gly Thr Gln His Gly Ile Arg Leu  
 1 5  
  
 <210> 38  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence:pro-memapsin 2  
 hydrolytic site #3 (aa 98-105 of SEQ ID NO:3)  
  
 <400> 38  
 Ser Ser Asn Phe Ala Val Gly Ala  
 1 5



<210> 39  
<211> 8  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence:pro-memapsin 2  
hydrolytic site #4 (aa 183-190 of SEQ ID NO:3)  
  
<400> 39  
Gly Leu Ala Tyr Ala Glu Ile Ala  
1 5